IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 11, No. 4, December 2022, pp. 1570~1578
ISSN: 2252-8938, DOI: 10.11591/ijai.v11.i3.pp1570-1578

Simulated trial and error experiments on productivity

Karn Thamprasert, Ahmad Yahya Dawod, Nopasit Chakpitak
International College of Digital Innovation, Chiang Mai University, Chiang Mai, Thailand

Article Info

Article history:

Received Sep 26, 2021
Revised Jun 4, 2022
Accepted Jul 3, 2022

Keywords:

Data-driven economic policy
Machine learning
Productivity
Simulated trial and error
Virtual economic environment

ABSTRACT

Trial and error experiments in socioeconomics were proved to be beneficial by Nobel prize laureates. However, replication is challenging and costly in term of time and money. The approach required interventions on human society, and moral issues have to be carefully considered in research designs. This work tried to make the approach more feasible by developing virtual economic environment to allow simulated trial and error experiments to take place. This research demonstrated the framework using 19 macroeconomic indicators in 6 interested categories to study the effect on productivity if each indicator value grew by 5 percent for each of 65 countries. Seven predictive models including some machine learning (ML) models were compared. Neural network dominated in accurateness and was selected as the core of the simulator. Experimented results are in full of surprises, and the framework acted as expected to be a data-driven guide toward country-specific policy making.

Corresponding Author:

Ahmad Yahya Dawod
International College of Digital Innovation, Chiang Mai University
239 Nimmanhaemin Road, Suthep, Muang, Chiang Mai, Thailand
Email: ahmadyahyadawod.a@cmu.ac.th

1. INTRODUCTION

In general, national socioeconomic policies have been made by domain professionals, experienced economists. Said experts are responsible in strategic planning for the whole country, faced with social pressure from expectations of the mass to solve crucial issues. There are concerns in effectiveness of the approach. First of all, experts are human, and human decisions and judgment differs depends on personal perspective rooted from personal experience [1]. In the darwinism viewpoint, human as an organism thrives to survive more than anything else [2]. Creating best possible policies to maximize public interests might or might not align with survival. Survival in modern settings includes conflicts of interests, hidden agendas, and so on. Despite leading companies and organizations had been working heavily in extracting information from big data, they rarely used those insights gained to make final decisions if the insights did not support their original belief [3]. Apart from ill-motivations and egoistic push, it is unsure that human brain alone is capable for such complex job.

Traditional approach in handling socioeconomic problems is to design the smartest solution for respective situation with experience and data available, then implement. If considered carefully, these smart solutions are equivalent to hypotheses in scientific process where claiming untested hypotheses as an answer would be risky and unacceptable [4]. While trial-and-error approach is proven to be superior in many ways [5], there are solid explanations underlying why it is underused in human-related decisions. While Banerjee *et al.* had been conducting more than 200 trial-and-error experiments to alleviate poverty worldwide [6] and get a Nobel prize in 2019, it is very hard to replicate and apply for more urgent problems as most of the experiments took 10-20 years to complete. The approach is way costlier than the intelligent designs, and there are ethical issues to be concerned and handled. This research aims to harness the powerfulness of the state-of-the-art trial-

and-error experiments approach while making the approach more feasible for social science studies to serve as a tool to suggest socioeconomic policies. With the help from machine learning (ML), this paper can build a virtual economic environment to be a safe experimenting place without involving real human.

This paper then simulates trial and error experiments of possible effects on national productivity in the virtual environment for each country when there are changes in labor force education structure, education outcome, education input, infrastructure, social environment, and technology creation. In the last part, we discuss important findings revealed from the model and give some examples of country-specific policy recommendations. Our virtual experiment approach is not expected be comparable to real experiments in accurateness, but should obviously serve as a cheaper, quicker, and safer alternative. ML popularity must be partly from its ability to perform both regression tasks, and classification tasks with elevated accuracy [7]. Given that productivity decomposition to construct the virtual economic environment should be counted as a regression problem, assumptions are that it will help with constraints of multicollinearity issues when more variables are employed in parametric models [8]. As most ML models are not restricted by linear specifications [9], accurateness is on the fly with the cost of harder interpretability and computation power [10]. With growing applications to predict in construction labor productivity [11], real estate [12], medical [13], epidermiology [14], finance [15], road maintenance [16], and other more. Random forest, gradient boosting, decision tree and neural network based models were proved to be outstanding choice for accuracy.

This work will compare as mentioned models, and some benchmarking models namely, ridge regression, and linear regression. Being productive suggests the ability to get more bang for the buck. Productivity measurement is intuitively a measure of economic efficiency if it is the output divided by total input. Given all inputs and outputs are in the same unit, then the productivity would simply be a meaningful multiplier. In macroeconomics studies, purchasing power parity (PPP) adjusted gross domestic product (GDP) represents total final goods in monetary unit to fend of price effect from inflation or deflation, often depicted in US Dollars [17]. But in the input counterpart is intricate as making goods involves multiple production factors, and most ways to mathematically combine input resources into an output goods makes the end result, total factor productivity (TFP) to lose its sense of being a pure multiplier of input to output. Consequently, the quantity TFP possess no intrinsic meaning. Though there are high volume of investigations on how to calculate TFP, yet strong conclusions do not exist [18]. Reported TFP are largely varies among publications due to diverse methodologies, data robustness, data frequency, data availability, and so on [19]. At this stage, choosing TFP is quite subjective depending on purpose. There were attempts to decompose productivity for underlying reasons behind efficient development. Assaf and Tsionas [20] has estimated and decomposed productivity in of tourism industry for 101 countries. Giving insightful findings for tourism destinations to tailor their strategies toward performing factors.

The work used Bayesian econometric approach to study source of temporal changes in productivity. The factors of interest are based on his previous work [21] namely, infrastructure quality indicators, human resource indicators, natural and environment quality indicators. Later in 2020, Maneejuk and Yamaka [22] studied and showed evidence of nonlinear impact on economic growth induced from both telecommunications related infrastructure and innovation. In their growth model [23] explicitly suggests education, research, and innovation aggregately as human capital. From these surveys, we decided to include some established macroeconomic indicators used widely in research, adjust, and reclassify into six mentioned categories namely, labor force education structure, education outcome, education input, infrastructure, social environment, and technology creation.

2. METHOD

In order to achieve the expected goals, the proposed framework comprised of five main steps. The first step is to clearly define objective variable which in this case is productivity of nations and factor variables of interest that will be tested to determine the effect on the objective variable if changed. Productivity was stated to be subjective so we will investigate further in this step. While factor variables were loosely defined earlier into six categories, will be specifically handpicked in the second step, data collection. Third is data transformation to make collected data suitable to use in training models. Fourth is to build the virtual economic environment with predictive models. Lastly, we simulate trial-and-error experiments by changing interested variable in the virtual environment and observe the outcomes.

2.1. Defining productivity

As stated earlier that total factor productivity choosing is subjective and depends on the purpose. Then in this step, we test some methods to calculate TFP in search for the one that suits our purpose best. There are three major approaches to estimate TFP, namely the production function approach, growth accounting approach, and non-parametric approach. Production functions map inputs to output [24]. Growth accounting approach is a technique centered around the growth of total factor productivity or rate of change

between years rather than its static value each year [25]. Lastly, non-parametric approach compares feasible input and output combinations based on the available data [26]. The productivity term is explicitly emitted in production functions while in growth accounting and non-parametric counterparts are not mentioned. The growth accounting focuses on rate of change across the timeline and the non-parametric emphasizes on how the input should combine. Then, production function is a solid choice to calculate TFP for this research.

The term productivity in production functions is also often used interchangeably with the technology factor [27]. As the approach is based on the concept of output is productivity or technology times combination of input factors. The plainest production function involves two input factors, capital (K) and labor (L) The production function of output Y with a level of productivity A as [28]:

$$Y = Af(K, L) \tag{1}$$

An industry standard production function is constant elasticity of substitution (CES) [29], the function is depicted as:

$$Y = A(\alpha K^{\rho} + (1-\alpha)L^{\rho})^{\frac{v}{\rho}} \tag{2}$$

where Y is output, A is productivity, K is capital, L is labor, α is share of capital in production, 1- α is share of labor in production, ρ is substitution parameter, v is degree of homogeneity. If $v = 1$ the production function is specified as constant return to scale, $v > 1$ for increasing return to scale, and $v < 1$ for decreasing return to scale [30]. CES is popular because of its flexibility as it exhibits the elasticity of substitution between factors, it could depict Cobb-Douglas, Leontief, and linear production in specific cases of its ρ value. If ρ is approaching 0 in the limit, then this CES would replicate Cobb-Douglas function:

$$Y = \lim_{\rho \to 0} A(\alpha K^{\rho} + (1-\alpha)L^{\rho})^{\frac{v}{\rho}} = AK^{\alpha}L^{1-\alpha} \tag{3}$$

If ρ is approaching 1 in the limit, then this CES would be a perfect linear substitute function:

$$Y = \lim_{\rho \to 1} A(\alpha K^{\rho} + (1-\alpha)L^{\rho})^{\frac{v}{\rho}} = A(\alpha K + (1-\alpha)L) \tag{4}$$

If ρ is approaching minus infinity, then this CES allows no substitution as similar in Leontief production function which specified fixed proportion of inputs:

$$Y = \lim_{\rho \to -\infty} A(\alpha K^{\rho} + (1-\alpha)L^{\rho})^{\frac{v}{\rho}} = A(Min(aK, bL)) \tag{5}$$

where a and b are pre-determined factor constant for capital and labor, respectively. The difference is if we use Leontief production function, productivity is fixated in a and b as fixed, making total productivity unextractable. While achievable in CES production function as productivity A is specified independently.

According to the CES in (2), parameters v and ρ are needed to decide optimal specifications, Alatas conducted a study in elasticity of substitution (ES) to examine if the value differs across countries [31]. The work modified Solow's CES to derive the growth regressions with the first-order Taylor expansion and assess ES depicted by σ for each country group. Groups are classified with the data-driven algorithm proposed by Phillips and Sul [32] using both panel and cross-sectional data on time-varying behavior of income per worker. Group 1 represents relatively more developed countries in term of income, while group 3 are least developed. The work's massive Taylor expansion yet naturally straightforward proved that ES varies across three country groups. The evidence of non-unitary ES suggests against using Cobb-Douglas's specification. Additionally, recent studies [33], [34] attack its ability to model modern world economy correctly. Alatas suggests σ=1.002 in group 1, σ=0.810 in group 2, σ=0.710 in group 3. We took the numbers as granted as well as the degree of homogeniety v=1 used in the work [31] to calculate our substitution parameter ρ. Noted that substitution parameter ρ could be calculated from the relation $\rho = \frac{\sigma - 1}{\sigma}$.

We computed the CES with granted parameters for two factors with the data both retrieved or derived from Penn's World Table 10.0 [35] spanning from 1970 up until 2019. Then, compare the value with the most basic means to calculate productivity, the single factor labor productivity. While being simple as real GDP at current PPP divided by hours worked per year, it comes with a meaningful unit of being output per hour worked.

The exploratory result shows declining productivity for most country as shown in Table 1 which contradicts astronomical technology progress in the last 50 years. It is observable that China have been

rapidly growing in both economy and technology in previous decades, yet productivity derived from CES shrinked by more than half. While labor productivity was more accurate to reflect its progress. In Thailand's case, CES productivity was cut by half from 1990 through 2000 when there was a financial crisis taking place which largely affect capital valuation. As well as in 2010 for the US which was recovering from 2008 subprime crisis. In the other hand, labor productivity could capture more in term of the ability to produce out of factor investment. CES model seems to be more prone to asset valuation, might be more useful in showing factor substitution, but clearly not fit for finding productivity. According to the exploratory findings, we choose labor productivity over the CES to represent productivity, the objective variable in our framework.

Table 1. Comparison between 2-factor CES productivity and labor productivity

	China		France		Thailand		United States	
Year	2-factor CES	labor only	2-factor CES	labor only	2-factor CES	labor only	2-factor CES	labor only
1970	1.221	1.540	0.629	21.238	1.085	2.458	0.705	33.538
1980	1.108	1.811	0.670	32.916	1.200	3.156	0.666	39.537
1990	1.028	2.243	0.585	39.459	1.113	4.457	0.680	45.761
2000	0.899	3.587	0.723	52.723	0.549	6.108	0.782	55.351
2010	0.662	8.305	0.475	61.947	0.671	10.329	0.663	68.816
2019	0.553	11.611	0.434	68.706	0.693	15.174	0.733	73.594

2.2. Data collection

In this step, we aim to collect as much data as we can to ease the training process. However, in most studies that work with macroeconomic indicators, one inevitable issue is missing data. Many unsatisfied data have to be filtered out. As in calculating labor productivity earlier which requires output-side real GDP at current PPPs (cgdpo), number of persons employed (emp), and average annual hours worked by worker (avh), there are only 65 countries providing sufficient data, limiting this research from studying worldwide productivity. Despite it seems that 65 countries are less than half of the world, these countries are responsible for 88.57% of global production in 2019. This work continues to collect factor variables in the same timeframe to associate the objective part for all six interested categories as followed.

Labor force education, we collect labor force with basic education, labor force with intermediate education, and labor force with advance education from the International Labour Organization labeled with lab_basic, lab_int, lab_adv respectively. All of these indicators represent in percentage of worker. Education outcome, we use mean Programme for International Student Assessment (PISA) score in mathematics, reading, and science from the Organisation for Economic Co-operation and Development (OECD) as pisa_math, pisa_read, and pisa_science. Education input, we retrieve pupil-teacher ratio in primary, secondary, and tertiary education level numbers from the UNESCO Institute of Statistics and define as ptr_basic, ptr_int, and ptr_adv.

Infrastructure, we grant overall logistics performance index (ranged from 1 to 5) from the world bank's logistic performance index surveys, percentage of individuals using the internet from the International Telecomunication Union, and the percentage of the population who have access to electricity from the United Nations Statistics Division. Variables are defined as inf_logistic, inf_net, inf_elec respectively. Social environment, we get life expectancy numbers in years (soc_lifeexp), adolescent aged between 15 and 19 fertility rates in birth per thousand (soc_adolefert), and percentage of urban population (soc_urbpop) from the United Nations Population Division.

Technology creation, we find numbers of researchers in R&D from the UNESCO Institute for statistucs, scientific and technical journal articles from National Science Foundation to be useful. Also, with the volume of industrial design applications and volume of patent applications from the World Intellectual Property Organization. These indicators are abbreviated as rnd_rs, rnd_jn, cre_indesign, and cre_patent.

As said, it is common to see missing data, but some could be saved. Linear interpolation is applied if applicable. There are some special cases. As PISA scores are missing in a few countries, we replace the blank with the minimum value for PISA scores of that subject. Missing data in researchers count, journal articles, industrial designs, and patents, we replaced blank with zero. In other edge cases, we replace not a number (NaN) by global average.

2.3. Data transformation

After filling blanks, we evaluate data distribution and correlations. It is also very common to see big difference among peers in income, GDP, and also productivity. These extremes are usually spotted by super positive skewness in data distribution and relieved by log-transformation. Taking natural log of non-negative variable added by one. Adding one prevent a case of $\ln(0) = -\infty$ from happening to ensure that transformed value remains non-negative.

In this research, we log-transformed labor force education indicators, individuals using internet, adolescent fertility rate, and technology creation indicators. There is one alienated variable: access to electricity (inf_elec), it is very negatively skewed because in most countries have almost 100% of population access to electricity. In prior to log-transform, we derive a new variable which stands for no access to electricity percent (inf_noelec = 1-inf_elec) to better separate sticky observations.

Next step is normalization. ML algorithm tend to converge faster and perform better when features are on a smaller scale. Also, it posed no harm to normalize. Consequently, it is a common practice to normalize the data before training ML models.

2.4. Building virtual economic environment

As this research intended to conduct panel experiments of all factors for all countries. Predictive models will be utilized as our laboratory, the virtual environment. It is concerned that high accuracy models are preferred without being overfitted. To prevent overfitting, this research split 80% of samples for training and the other 20% for validation. Then, fit prepared data in the proposed models, specifications are as followed.

In random forest regression, gradient boosting regression, decision tree regression, polynomial regression, ridge regression, and linear regression modelling, multiple packages in python library sklearn is utilized. It is known that random forest and alike models will stabilize after some degree of trees (n_estimators), we use stepwise refinement to decide where to stop. For learning rate tuning in the gradient boosted model, it is grid searched along with said hyperparameter from 0.001, 0.003, 0.01, 0.03, 0.1, 0.3, up until 1, multiplies by an approximate of 3 at a time. For alpha value in ridge regression, repeated K-fold is used to evaluate the value which returns lowest negative mean absolute error from the alpha range of 0 to 100, with 0.1 increments rate.

Artificial neural network (ANN) is modelled with keras library. The architecture consists of 20 units in the input layer, one hidden layer with 256 units using rectified linear unit (ReLU) activation function, and one output units. 20 input units take 19 features plus an array of ones for intercept. In search for the best candidate, criteria are mean-squared-errors (MSE) and R-squared. The one with least MSE and highest R-squared will be selected.

2.5. Simulate trial and error experiments

Trial and error experiments use real feedback loop to study relationship between actions and reactions. It was mentioned that physical trials are ideal, but in many cases, the cost are too high to conduct the test. In the previous step, we built and chose the best predictive models to be virtual environment for this study. Tracking parameters can be intimidating, and researchers might want to explore and reverse engineer to find profounding universal law by digging into the black box and, but those mechanisms are not concerned in this research. Instead, we found it more useful not to bother and treat it as a playground. Recalling that we know that the box would be in this form regardless of model selections.

$$\hat{Y}_{c,t} = F(1, X^1_{c,t}, X^2_{c,t}, \dots, X^k_{c,t}) \quad (6)$$

where $\hat{Y}_{c,t}$ stands for predicted productivity of country c at time t, $F(\bullet)$ is unknown function, $X^i_{c,t}$ represents factor i of country c at time t. An interesting sample question to prove this framework useful could be "What if country c has enough capacity to improve only one factor from 19 factors by δ percent, what should the country decide to reach best possible productivity?" In ceteris paribas, other things stay still, improving one indicator $X^2_{c,t}$ by $\delta \in \mathbb{R}$ percent can be depicted as:

$$\tilde{Y}^2_{c,t} = F(1, X^1_{c,t}, (1 + \frac{\delta}{100})X^2_{c,t}, \dots, X^k_{c,t}) \quad (7)$$

Then the evaluation metric of the experiment effect will essentially be $\varphi^2_{c,t} = \frac{\tilde{Y}^2_{c,t} - \hat{Y}_{c,t}}{\hat{Y}_{c,t}}$. Repeat doing this way with the same δ in the same study timeframe t for every country and X factors should give a panel of country-specific comparable results based on their current stage of development. This research will give an example of $\delta = 5, t = 2019$, panel testing every country on the change of every factor to show the simulated trial and error result and discuss its findings along with possible policy recommendations.

3. RESULTS AND DISCUSSION

Proceeded through steps in the methodology. Data is collected, clean, and transformed as stated. In building predictive models, ANN outperformed other models in predictability upon validation with the lowest MSE of 0.004 and highest R-squared of 99.54%. Although random forest regression performed as

nearly as close to the leader with MSE of 0.0053 and R-squared of 99.4%. Details are shown in the Table 2. Noted that models with linear restriction such as linear regression and ridge regression performed significantly lower than those unrestricted. These findings also show that in labor productivity fitting, non-linear models are preferred if the goal is to maximize accuracy as in this work.

Table 2. Model fitting result

Algorithm	MSE	R-squared
Random Forest Regression	0.0053	0.9940
Gradient Boosting Regression	0.0182	0.9793
Artificial Neural Network	0.0040	0.9954
Decision Tree Regression	0.0124	0.9859
Ridge Regression	0.0847	0.9038
Second-degree Polynomial Regression	0.0184	0.9791
Linear Regression	0.0847	0.9038

The result suggests ANN to be the core of the simulation. Testing our work by giving an increase of 5% ($\delta = 5$) for each indicator for each country, small part of the results can be found in Table 3. Condensed presentations are due to space limitation. Countries represented in the table are selected to represent how distinct cultures from different income levels react to simulated changes. Each percentage numbers in the table represents the effect of each experiment $\varphi^{i}_{c,2019}$.

Table 3. Percentage effect on 2019 labor productivity if corresponding indicator improved by 5% (δ=5)

	Bangladesh	China	Germany	Hong Kong SAR	South Africa	Thailand	United States	World
lab_basic	-0.76%	2.20%	0.63%	3.04%	-1.20%	0.85%	0.26%	0.46%
lab_int	-1.23%	-5.34%	1.80%	-4.02%	0.51%	-1.52%	-1.73%	-1.21%
lab_adv	2.04%	0.28%	0.15%	2.32%	2.06%	1.48%	-0.51%	0.40%
pisa_math	-0.33%	-1.26%	-0.42%	2.53%	0.35%	-0.07%	-1.21%	-0.60%
pisa_read	-2.26%	-1.82%	-1.34%	4.84%	-0.41%	-2.54%	1.96%	0.83%
pisa_science	4.50%	6.86%	1.77%	-1.03%	1.29%	2.97%	3.60%	1.89%
ptr_basic	4.12%	0.63%	-0.30%	-11.27%	5.51%	4.43%	0.26%	0.39%
ptr_int	3.04%	0.53%	0.43%	1.00%	-2.34%	1.59%	0.25%	0.16%
ptr_adv	-1.80%	-2.29%	-0.63%	4.96%	-2.10%	-2.39%	0.50%	-0.08%
inf_logistic	4.14%	-1.34%	2.97%	11.61%	2.86%	1.67%	0.06%	2.02%
inf_net	1.23%	2.38%	3.85%	4.68%	0.56%	-2.75%	3.21%	3.35%
inf_noelec	-2.78%	0.00%	0.00%	0.00%	-1.37%	-0.05%	0.00%	-0.31%
soc_lifeexp	-5.42%	-4.34%	-9.13%	8.59%	-2.04%	-10.48%	-5.86%	-10.88%
soc_adolefert	1.19%	2.24%	0.14%	0.88%	2.33%	1.45%	1.97%	1.60%
soc_urbpop	-0.77%	0.46%	0.97%	-3.38%	-1.87%	-0.65%	0.20%	0.76%
rnd_rspop	-1.14%	1.07%	-1.99%	0.00%	-1.30%	-1.60%	1.62%	-0.02%
rnd_jnpop	-1.53%	-1.36%	-2.95%	0.00%	-0.86%	-1.75%	0.08%	-0.76%
cre_indesign	0.78%	0.49%	1.26%	1.80%	0.78%	0.75%	0.52%	0.53%
cre_patent	-0.65%	2.52%	1.80%	-0.93%	-1.35%	0.43%	1.74%	0.49%

The result shows expected dissimilar effects on labor productivity given the same improvement across countries which could not be easily achieved in traditional parametric models. Followings are important findings from the result in the Table 3. According to our experiments, nudging the labor force education structure variables (lab_basic, lab_int, lab_adv) showed some notable findings. For the world in overall, increasing the proportion of workers with only basic education, and advance education by five percent improved labor productivity by 0.46% and 0.40% respectively. While applying the same intervention with high school graduates' proportion adversely affect productivity by -1.21%. This suggests workers to opt for advanced education or stay at the basic level. However, things are not going in the same way if we study in each country. In Bangladesh and South Africa, changing the structure by having 5% more worker with basic skill should expect slight decline in productivity. Bangladesh would need to go with full thrusters to promote advanced education aiming for 2.04% improve in productivity. But in South Africa, the situation is different. There are big opportunities in advanced education yet allowing more intermediate level workers contrary to Bangladesh. China in the other hand, is expected to gain 2.2% for more worker with just basic education. The situation got more extreme in China for high school graduates. Adding 5% more expected a big decline of 5.34%. Plausible policy for China might be encouraging medium skill labor to work abroad while importing basic skill labor. These findings in this category alone proved the framework to be useful works specifically to each nations' context.

In education outcome where PISA scores are considered, most countries are already good enough at math and reading, while there is a significant room to improve in science especially in Bangladesh, China,

Thailand, and United States. Things are different in Hong Kong; they are better in science but not in reading and math. For education input considering pupil to teacher ratio in three levels, more teachers for primary and secondary school, with less teachers in tertiary institutes are suggested in Thailand and Bangladesh. Again, Hong Kong is in the opposite side. Five percent more teachers in primary level is expected to cause a -11.27% decline in productivity while more professors are much needed.

Out of all choices, improving internet accessibility will boost productivity the best for the world for a big gain of 3.35%. This promotion could be applied to most places and wait for productivity gain. The only exception is Thailand which is quite saturated in internet accessibility [36]. More internet addiction in Thailand will lead to a -2.75% decline in productivity. Second best stimulus intervention for the world efficiency is to improve logistics infrastructure with expected 2.02% gain. This suggestion stands for everywhere but China. As noticeable that China has been building a lot in the last few decades, building more at this moment might be too early and expected to cause a negative effect. At the time of writing (2022), there is an oversupply in real estates led by rapid urban expansion in China that might soon result in another financial crisis [37]. Maybe the framework that is using 2019 data might be giving a clue before it happens.

If ones give a glance, electricity did not play a big role here. It is from the fact that almost everywhere on earth could already access to electricity, making majority of its effect seems irrelevant. There's a case in Bangladesh though, if there are 5 percent more people without electricity, then -2.78% in productivity is predicted. Correct interpretation is crucial here. Electricity is vital, but the way this framework worked, adding five percent to the current people without electricity for most countries with already low amount of these people will be small. Then, the effect should be miniscule in overall. The issue emphasizes that this framework can not be used to judge the effect of electricity as well as other factors on productivity, it is designed to answer "what-if" questions on variants from its current state instead.

Social environment factors of our interests are life expectancy (soc_lifeexp), adolescent fertility rate (soc_adolefert), and urban population (soc_urbpop). The result showed unexpected strong decline in productivity if current life expectancy increased by 5% around the world except in Hong Kong area which confirmed positive effect. A proper explanation would be if life expectancy is high, there will be more retired elders which discontinued to actively contribute for the economy. The exception could be due to Hong Kong's high official retirement age of 65 and seniors are in favor of working than in other cultures [38]. Another shocking finding is that the algorithm promotes adolescent fertility across the globe. The model might have some points in it toward productivity. Increasing urban population outcomes are varying but there is a pattern. In overall, the framework suggests people to move into urban area except in places where the cities are already too dense.

These social factors give some examples of morally conflict cases. Raising life expectancy is a noble yet complex thing to do, but lowering it is unethical. It is also not a good idea to encourage adolescent fertilization and do real trial and error experiments for many reasons. This virtual economy is a safer place to do such experiments. However, no one should ever make a policy to lower life expectancy or promote adolescent fertility even the model says so. The model just did its naïve work to optimize labor productivity without caring other things.

Technology creation factors exhibit interesting findings. Having more researchers lead to a decline in productivity in most countries especially in Germany (-1.99%) with only two exceptions, United States and China, which shows positive effect. Intuitively, places with more established researchers have already surpassed the unyielding investment phase while others have to invest much more to catch up. Moreover, the evidence showed scientific publications as an ineffective way to promote labor productivity for all countries. It might be due to high financial, time, and opportunity cost while taking ages to be implemented. In the other hand, industrial designs which take less time to apply rewarded more while patents are in between.

4. CONCLUSION

This research built a virtual economic environment as a playground to simulate trial and error experiments with neural networks at its core. First purpose and contribution of the framework is to mitigate moral and sensitive issues in the real human-related trial and error experiments. Moreover, instead of having to invest decades and money, ones can safely simulate desired action in the virtual environment if it is accurate enough. We demonstrated by using 19 macroeconomic indicators in 6 categories to study the effect on labor productivity. Apart from other works of its kind, the framework did not expect to crack what builds up in countries productivity by studying parameters and stereotyping the whole world. Instead, our work built on difference of countries. The results are proved to be country-specific guide toward personalized policy recommendations as proposed. The work compared 7 models and neural networks proved to be the best in accurateness with R-squared of 0.9954. High accuracy came from allowance of interactions between variables and absence of linear assumptions. Negative issue to expect from neural networks is

interpretability. Many other works tried to dig inside the model to give a closed form explanation. We prefer to focus our effort in doing trial and error experiments from its prediction. That is why we called our work a simulation rather than a model. Answering to the productivity question, we shocked each factor of interest by five percent for each country at a time forming a panel tests and observed the changes. Experimented results confirmed disparity across countries. Worked solution in one place failed in another place. The framework proved to be useful as intended. Simulation revealed numerous unexpected insights detailed in previous section. There is even a possibility that it might be able to spot an economic crisis before it happens. However, the framework's capability might be confusing. To clarify, the framework is designed to be used in the "what-if" scenario rather than proving relationship or correlations between x and y. In a part of the result, we experimented with some sensitive social factors effect on productivity. Our virtual economy acted as a safe place to do these intimidating trials where conducting the same test in real life could take years, cause a fortune, harder to extract the result, and secure a title of public enemy if not turned down by ethics committee. In overall, the research goals are satisfied. The framework could be applied with other research interests especially in socioeconomics. There are plenty of room to extend and improve this to be a practical data-driven policy making tool. Though, it should be use with precautions as the tool focused on objective optimization without caring anything else.

ACKNOWLEDGEMENTS

We would like to express our thanks to reviewers for reading this paper, to Graduate School, Chiang Mai University for funding through Junior Research Grant number JRCMU2564_072 , and to International College of Digital Innovation, CMU for academic supports.

REFERENCES

[1] F. Camastra *et al*., "A fuzzy decision system for genetically modified plant environmental risk assessment using Mamdani inference," *Expert Systems with Applications*, vol. 42, no. 3, pp. 1710–1716, Feb. 2015, doi: 10.1016/j.eswa.2014.09.041.
[2] C. Darwin, *On the Origin of Species*. London: Penguin Classics, 2009.
[3] S. D. Levitt and S. J. Dubner, *Think like a freak*. Harper Audio, 2014.
[4] T. Harford, *Adapt: Why success always starts with failure*. Seoul: Farrar, Straus and Giroux, 2011.
[5] M. Ridley, "Economics: Trial and error," *Nature*, vol. 474, no. 7349, pp. 32–33, Jun. 2011, doi: 10.1038/474032a.
[6] A. V. Banerjee, E. Duflo, and M. Kremer, "The influence of randomized controlled trials on development economics research and on development policy," in *The State of Economics, the State of the World*, The MIT Press, 2020, pp. 482–488.
[7] S. García, A. Fernández, J. Luengo, and F. Herrera, "A study of statistical techniques and performance measures for genetics-based machine learning: accuracy and interpretability," *Soft Computing*, vol. 13, no. 10, pp. 959–977, Dec. 2009, doi: 10.1007/s00500-008-0392-y.
[8] A. Alin, "Multicollinearity," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, no. 3, pp. 370–374, Mar. 2010, doi: 10.1002/wics.84.
[9] H. Liu *et al*., "A nonlinear regression application via machine learning techniques for geomagnetic data reconstruction processing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 1, pp. 128–140, Jul. 2018, doi: 10.1109/tgrs.2018.2852632.
[10] D. V Carvalho, E. M. Pereira, and J. S. Cardoso, "Machine learning interpretability: A survey on methods and metrics," *Electronics*, vol. 8, no. 8, p. 832, Jul. 2019, doi: 10.3390/electronics8080832.
[11] S. Golnaraghi, O. Moselhi, S. Alkass, and Z. Zangenehmadar, "Predicting construction labor productivity using lower upper decomposition radial base function neural network," *Engineering Reports*, vol. 2, no. 2, p. e12107, Jan. 2020, doi: 10.1002/eng2.12107.
[12] T. Mohd, S. Jamil, and S. Masrom, "Machine learning building price prediction with green building determinant," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 9, no. 3, p. 379, Sep. 2020, doi: 10.11591/ijai.v9.i3.pp379-386.
[13] N. A. M. Nor, A. Mohamed, and S. Mutalib, "Prevalence of hypertension: predictive analytics review," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 9, no. 4, p. 576, Dec. 2020, doi: 10.11591/ijai.v9.i4.pp576-583.
[14] K. Mahmoud *et al*., "Prediction of the effects of environmental factors towards COVID-19 outbreak using AI-based models," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 10, no. 1, p. 35, Mar. 2021, doi: 10.11591/ijai.v10.i1.pp35-42.
[15] M. R. Pahlawan, E. Riksakomara, R. Tyasnurita, A. Muklason, F. Mahananto, and R. A. Vinarti, "Stock price forecast of macro-economic factor using recurrent neural network," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 10, no. 1, p. 74, Mar. 2021, doi: 10.11591/ijai.v10.i1.pp74-83.
[16] S. M. Piryonesi and T. E. El-Diraby, "Data analytics in asset management: Cost-effective prediction of the pavement condition index," *Journal of Infrastructure Systems*, vol. 26, no. 1, p. 4019036, Mar. 2020, doi: 10.1061/(asce)is.1943-555x.0000512.
[17] K. Fukao, D. Ma, and T. Yuan, "Real GDP in pre-war East Asia: A 1934–36 benchmark purchasing power parity comparison with the US," *Review of Income and Wealth*, vol. 53, no. 3, pp. 503–537, Sep. 2007, doi: 10.1111/j.1475-4991.2007.00243.x.
[18] M. G. Tsionas and M. L. Polemis, "On the estimation of total factor productivity: A novel Bayesian non-parametric approach," *European Journal of Operational Research*, vol. 277, no. 3, pp. 886–902, Sep. 2019, doi: 10.1016/j.ejor.2019.03.035.
[19] J. Van Biesebroeck, "The sensitivity of productivity estimates: Revisiting three important debates," *Journal of Business & Economic Statistics*, vol. 26, no. 3, pp. 311–328, Jul. 2008, doi: 10.1198/073500107000000089.
[20] A. G. Assaf and M. Tsionas, "The estimation and decomposition of tourism productivity," *Tourism Management*, vol. 65, pp. 131–142, Apr. 2018, doi: 10.1016/j.tourman.2017.09.004.
[21] A. G. Assaf and E. G. Tsionas, "Incorporating destination quality into the measurement of tourism performance: A Bayesian approach," *Tourism Management*, vol. 49, pp. 58–71, Aug. 2015, doi: 10.1016/j.tourman.2015.02.003.
[22] P. Maneejuk and W. Yamaka, "An analysis of the impacts of telecommunications technology and innovation on economic growth," *Telecommunications Policy*, vol. 44, no. 10, p. 102038, Nov. 2020, doi: 10.1016/j.telpol.2020.102038.
[23] R. E. Lucas, "On the mechanics of economic development," *Journal of Monetary Economics*, vol. 22, no. 1, pp. 3–42, Jul. 1988,

doi: 10.1016/0304-3932(88)90168-7.
[24] R. M. Solow, “The production function and the theory of capital,” *The Review of Economic Studies*, vol. 23, no. 2, pp. 101–108, 1955, doi: 10.2307/2296293.
[25] N. CRAFTS, “Productivity growth in the industrial revolution: A new growth accounting perspective,” *The Journal of Economic History*, vol. 64, no. 02, pp. 521–535, Jun. 2004, doi: 10.1017/s0022050704002785.
[26] A. Hailu and T. S. Veeman, “Non-parametric productivity analysis with undesirable outputs: an application to the Canadian pulp and paper industry,” *American Journal of Agricultural Economics*, vol. 83, no. 3, pp. 605–616, Aug. 2001, doi: 10.1111/0002-9092.00181.
[27] K. I. Carlaw and R. G. Lipsey, “Productivity, technology and economic growth: what is the relationship?,” *Journal of Economic Surveys*, vol. 17, no. 3, pp. 457–495, Jul. 2003, doi: 10.1111/1467-6419.00201.
[28] A. Zellner, J. Kmenta, and J. Dreze, “Specification and estimation of Cobb-Douglas production function models,” *Econometrica*, vol. 34, no. 4, p. 784, Oct. 1966, doi: 10.2307/1910099.
[29] D. F. Heathfield, “The constant elasticity of substitution function,” in *Production Functions*, Macmillan Education {UK}, 1971, pp. 45–69.
[30] R. Klump, P. McAdam, and A. Willman, “The long-term sucCESs of the neoclassical growth model,” *Oxford Review of Economic Policy*, vol. 23, no. 1, pp. 94–114, Mar. 2007, doi: 10.1093/oxrep/grm003.
[31] S. Alatas, “Income and factor substitution: an investigation on the Solow growth model under the constant elasticity of substitution,” *Journal of Economic Studies*, vol. 49, no. 3, pp. 397–421, Mar. 2021, doi: 10.1108/jes-07-2020-0335.
[32] P. C. B. Phillips and D. Sul, “Transition modeling and econometric convergence tests,” *Econometrica*, vol. 75, no. 6, pp. 1771–1855, Nov. 2007, doi: 10.1111/j.1468-0262.2007.00811.x.
[33] S. Gechert, T. Havranek, Z. Irsova, and D. Kolcunova, “Measuring capital-labor substitution: The importance of method choices and publication bias,” *Review of Economic Dynamics*, vol. 45, pp. 55–82, Jul. 2022, doi: 10.1016/j.red.2021.05.003.
[34] S. Gechert, T. Havránek, Z. Havránková, and D. Kolcunova, “Death to the Cobb-Douglas production function,” Hans-Böckler-Stiftung, Institut für Makroökonomie und Konjunkturforschung (IMK), 2019. doi: 10.31222/osf.io/6um5g.
[35] R. C. Feenstra, R. Inklaar, and M. P. Timmer, “The next generation of the Penn World Table,” *American Economic Review*, vol. 105, no. 10, pp. 3150–3182, Oct. 2015, doi: 10.1257/aer.20130954.
[36] S. Kemp, “Digital 2022: Global overview report,” *DataReportal–Global Digital Insights*, 2022. https://datareportal.com/reports/digital-2022-global-overview-report (accessed Feb. 23, 2022).
[37] “China evergrande debt crisis: five developers on the brink,” *Asia Financial*, 2022. https://www.asiafinancial.com/china-evergrande-debt-crisis-five-developers-on-the-brink (accessed Feb. 23, 2022).
[38] S. Xinqi, “hy Hong Kong’s seniors are in favour of working in old age – they can’t survive any other way,” *South China Morning Post*, 2018. https://www.scmp.com/news/hong-kong/hong-kong-economy/article/2144048/why-hong-kongs-seniors-are-favour-working-old-age (accessed Feb. 23, 2022).

BIOGRAPHIES OF AUTHORS

Karn Thamprasert Received Bachelor and Master of Economics degree from Chiang Mai University, Thailand in 2009, and 2018 respectively. He was awarded with certificates of achievement in biochemistry, genetics, immunology, and physiology from Harvard Medical School. He also earned Micromasters of Statistics and Data Science credential from MITx. He is currently a Ph.D. candidate in Digital Innovation and Financial Technology in Chiang Mai University. He holds managing director position in Q Path Lab Company Limited, a medical software R&D firm and other 4 more companies. His research includes machine learning, economics, and international trade policy. He has one publication in international journals and conferences. He can be contacted at email: karn_thamprasert@cmu.ac.th.

Dr. Ahmad Yahya Dawod He is currently a lecturer at International College of Digital Innovation Chiang Mai University, Thailand. He received his Ph.D. degree in Machine Learning and Artificial Intelligence from the National University of Malaysia in 2018 with the topic “Hand gesture recognition based on isolated and continuous sign language”. He also graduated his master’s degree in computing and informatics from Multimedia University of Malaysia and had his bachelor’s degree in computer science from The University of Mustansirya of Iraq. His research includes machine learning, pattern recognition, computer vision, robotics, and artificial intelligence. He had published 19 articles up to date with more than hundred citations. He can be contacted at email: ahmadyahyadawod.a@cmu.ac.th.

Dr. Nopasit Chakpitak He is a researcher in artificial intelligence, knowledge engineering and management. His knowledge engineering application domains include power industry, asset management and SMEs clusters. He is the founder dean of the College of Arts, Media and Technology, Chiang Mai University. Now he is an assistant professor in knowledge management and software engineering. His consultancy works are mainly in knowledge management implementation for government agencies, state and private enterprises by using knowledge engineering methodology. He also involves in several European funded projects in ICT and it applications. His current works are on smart city including smart transportation, smart tourism, smart education, smart energy, etc. This includes relevance disruptive technologies: 4G/5G mobile applications, cloud computing, internet of things, big data, block chain, FinTech, artificial intelligence, cyber security, IP and security laws. He can be contacted at email: nopasit@cmuic.net.